
MarTREC UTC Data Management Plan

A Data Management Plan created using DMPTool

Creator: Heather Nachtmann

Affiliation: Non Partner Institution

Template: Department of Transportation (DOT)

Last modified: 06-13-2017

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customize it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

MarTREC UTC Data Management Plan

Data description

To comply with the U.S. Department of Transportation (USDOT) Public Access Plan (<https://www.transportation.gov/mission/open/official-dot-public-access-plan-v11>), the MarTREC UTC will require each individual researcher to submit detailed data descriptions for their individual research projects per this data management plan (DMP) as outlined in the guidance.

1. Name the data, data collection project, or data producing program.
2. Describe the purpose of the research.
3. Describe the data that will be generated in terms of nature and scale (e.g., numerical data, image data, text sequences, video, audio, database, modeling data, source code, etc.).
4. Describe methods for creating the data (e.g., simulated; observed; experimental; software; physical collections; sensors; satellite; enforcement activities; researcher-generated databases, tables, and/or spreadsheets; instrument generated digital data output such as images and video; etc).
5. Discuss the period of time data will be collected and frequency of update.
6. If using existing data, describe the relationship between the data you are collecting and existing data.
7. List potential users of the data.
8. Discuss the potential value of the data have over the long-term for not only your institution, but also for the public.
9. If you request permission not to make data publicly accessible, explain rationale for lack of public access.
10. Indicate the party responsible for managing the data.
11. Describe how you will check for adherence to this data management plan.

Data format and metadata standards

To the maximum extent practicable, MarTREC investigators will use platform-independent and non-proprietary formats to ensure maximum utility of the data in the future.

1. All investigators will be required to have all non-proprietary final datasets in the standard data format of the field such as csv.
2. If this is not possible, investigators will be required to describe how they will document the alternative formats they are using and why the file format(s) they are using is(are) not able to be in the standard data format such as csv.
3. If investigators are using proprietary data formats, they will be required to discuss their rationale for using those standards and formats and receive prior approval of the MarTREC center director.
4. Investigators will be required to describe the data process log to clarify the final version of data shared to the public.
5. Investigators will list what documentation they will be creating in order to make the data understandable by other researchers.
6. Investigators will indicate what metadata schema they are using to describe the data. If the metadata schema is not one standard for their field, discuss their rationale for using that scheme.
7. Investigators will have to describe how the metadata will be managed and stored.
8. Investigators will indicate what tools or software is required to read or view the data.
9. Investigators will describe their quality control measures.

Policies for access and sharing

Investigators will be required to address any access restrictions in the project DMP they submit to the MarTREC UTC DMP. For project DMPs, investigators will address issues and outline the efforts they will take to provide informed consent statements to participants, the steps they will take to protect privacy and confidentiality prior to archiving their data, and any additional concerns (e.g., embargo periods for their data). If necessary, they will describe any division of responsibilities for stewarding and protecting the data among other project staff. If investigators will not be able to deidentify the data in a manner that protects privacy and confidentiality while maintaining the utility of the dataset, investigators will describe the necessary restrictions on access and use. If an individual research project includes human subject research, investigators will be required to go through the University of Arkansas Institutional Research Board (IRB) or their home institutions' IRB, if they have one.

In general, investigators will be required to address the following in their project DMPs:

1. Describe what data will be shared, how data files will be shared, and how others will access them.
2. Indicate whether the data contain private or confidential information. If so,
 - Discuss how they will guard against disclosure of identities and/or confidential business information.
 - State the party responsible for protecting the data.
 - List what processes they will follow to provide informed consent to participants.
1. Describe what, if any, privacy, ethical, or confidentiality concerns are raised due to data sharing.
2. If applicable, describe how they will de-identify their data before sharing. If not,
 - Identify what restrictions on access and use they will place on the data.
 - Discuss additional steps, if any, they will use to protect privacy and confidentiality.

Policies for re-use, redistribution, derivatives

The University of Arkansas or the home institution of the investigators holds the IP for data created by the project. Investigators will be required to describe if they are transferring rights to the data archive. If they do not describe this, their home institution maintains the rights. Investigators will be required to cite the data source and license under which they used the data in their project DMPs.

In general, investigators will address the following in their project DMPs:

1. Name who has the right to manage the data.
2. Indicate who holds the intellectual property rights to the data.
3. List any copyrights to the data. If so, indicate who owns them.
4. Discuss any rights to be transferred to a data archive.
5. Describe how their data will be licensed for reuse, redistribution, and derivative products.

Plans for archiving and preservation

Plans for archiving will support the capture and provision of the U.S. Federal Government Project Open Data Metadata Schema. In addition, the archive will support the creation and maintenance of persistent identifiers (e.g., DOIs, handles, etc.) and will provide for maintenance of those identifiers throughout the preservation lifecycle of the data.

1. MarTREC will archive all publications and data on CERN's Zenodo, <https://zenodo.org/communities/martrec>, which is conformant with US DOT guidelines as described at <https://ntl.bts.gov/publicaccess/evaluatingrepositories.html>.
2. When a project submits a final report, the investigators will have 60 days to archive their data on Zenodo.
3. Investigators will maintain and back-up data until it is uploaded to Zenodo.
4. Zenodo's procedures and policies for back-up, data recovery, retention, security and integrity are outlined in <https://zenodo.org/policies>.
5. Zenodo provides how back-up, disaster recovery, off-site data storage, and other redundant storage strategies will be used to ensure the data's security and integrity.
6. Zenodo will retain data for the lifetime of the repository. This is currently the lifetime of the host laboratory CERN, which currently has an experimental program defined for the next 20 years at least.
7. Each data upload in Zenodo gets a Digital Object Identifier (DOI) to make them easily and uniquely citable.